

新颖性会话推荐机制研究

崔博伦, 孙丽梅, 曹剑钊

沈阳建筑大学计算机科学与工程学院, 辽宁 沈阳

收稿日期: 2026年1月5日; 录用日期: 2026年2月4日; 发布日期: 2026年2月12日

摘要

目前会话推荐模型大多通过计算项目物品与用户会话的准确性匹配度,从而决定用户的推荐列表。但是,仅依靠准确性匹配度进行推荐导致存在推荐新颖性不足、推荐结果倾向热门化问题。为解决上述问题,本文提出了一种能够提升会话推荐列表新颖性的新颖性机制,将其运用在会话推荐模型中以观察效果。本文首先将用户与会话推荐系统的交互行为数据构建为有向会话图,使用图神经网络的门控更新机制融合邻域信息与全局偏好,为每一个项目生成向量信息;然后,将会话的局部兴趣与全局偏好的注意力加权融合,计算会话的向量,并通过会话的向量与项目进行相似性计算,得出每个项目的推荐得分;最后,引入基于流行度的物品新颖度指标,可以通过修改参数 α 平衡新颖性与准确性,得出每个物品的综合得分,取出综合得分排名前列的物品作为推荐列表。在调整参数 α 的过程中,本文根据准确性和新颖性实验结果,将新颖化推荐分为三种模式:精准推荐模式、新颖性推荐模式、新颖性优先模式。系统管理者可以根据不同的场景灵活设置参数 α 以完善用户体验。本文为提升会话推荐系统的新颖性推荐能力提供了一种新的技术路径。

关键词

会话推荐系统, 图神经网络, 新颖性推荐

Novel Session-Based Recommendation Mechanism

Bolun Cui, Limei Sun, Jianzhao Cao

School of Computer Science and Engineering, Shenyang Jianzhu University, Shenyang Liaoning

Received: January 5, 2026; accepted: February 4, 2026; published: February 12, 2026

Abstract

Currently, most session-based recommendation models determine the user's recommendation list by calculating the accuracy matching degree between items and the user session. However, relying solely on accuracy matching for recommendations leads to issues such as insufficient

recommendation novelty and a tendency for results to favor popular items. To address these problems, this paper proposes a novel mechanism capable of enhancing the novelty of session-based recommendation lists and applies it to session-based recommendation models to observe its effects. This paper first constructs the interaction data between users and the session-based recommendation system as a directed session graph. It then employs the gated update mechanism of a graph neural network to fuse neighborhood information and global preferences, generating vector representations for each item. Next, the local interests of the session and the global preferences are fused via attention-weighted aggregation to compute the session vector. The similarity between the session vector and items is calculated to derive the recommendation score for each item. Finally, a popularity-based item novelty metric is introduced, which allows balancing novelty and accuracy by adjusting parameter x , resulting in a comprehensive score for each item. The top-ranked items based on this comprehensive score are selected as the recommendation list. During the process of adjusting parameter x , this paper categorizes novelty-enhanced recommendations into three modes based on experimental results of accuracy and novelty: Accuracy-First Mode, Novelty-First Mode, and Novelty-Priority Mode. System administrators can flexibly set parameter x according to different scenarios to improve the user experience. This paper provides a new technical approach for enhancing the novelty recommendation capability of session-based recommendation systems.

Keywords

Session-Based Recommendation Systems, Graph Neural Networks, Novel Recommendation

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着移动互联网应用迅速发展，海量数据的传播导致严重的信息过载问题。为了解决信息过载问题，推荐系统成为一种能够帮助用户快速、准确地找到自己需要信息的解决方案[1]。在解决信息过量和提升用户浏览体验这两方面，推荐系统发挥着不可或缺的作用。推荐系统收集并分析用户浏览历史、购买行为、兴趣偏好等类型的数据，利用复杂的算法与模型，为用户提供个性化的推荐内容。推荐系统能够从海量信息中筛选出与用户信息匹配的内容，帮助用户找到他们可能感兴趣，但又难以在众多信息中自行挖掘的有价值资源[2]，使得用户不需要在大量的信息中浪费大量时间和精力进行搜索与筛选，极大地节省了用户的时间、精力，有效提升了信息获取的效率[3]。

传统推荐系统模型依赖于长期的用户与系统的交互行为(例如评分、评论记录、点击记录、购买记录等)进行推荐列表生成。但是在实际场景中，用户很有可能会进行短期或者匿名的与推荐系统交互行为。这种场景突出了强实时性的特点，此时用户身份信息和长期偏好无法获取，仅可使用当前会话内的有限数据资源建立模型[4]。为解决上述问题，依据会话数据进行建模的会话推荐系统应运而生。在目前会话推荐系统的研究中，大部分研究仅以准确率为中心，导致用户视野受限、部分小众商品展示机会少、平台内容单一且重复度较高、资源信息分配不均衡等不利影响。

2. 问题定义

2.1. 会话推荐系统

集合 $U = \{u_1, u_2, \dots, u_{|U|}\}$ 代表了数据集中的所有的用户，集合 $V = \{v_1, v_2, \dots, v_{|V|}\}$ 代表了数据集中的所有

项目。某个用户在某一会话中交互的项目集合构成一个会话 $s_i = \{v_1, v_2, \dots, v_{|s|}\}$ ，数据集中所有的会话则构成了会话集合 $S = \{s_1, s_2, \dots, s_{|S|}\}$ 。可将会话推荐系统的功能定义为：根据一个给定的会话 $\{v_1, v_2, \dots, v_x\}$ ，预测出若干个可能与用户交互的项目 $v_{(x+1)}$ [5]。

2.2. 点击命中率

点击命中率($Hit @ N$)的定义为：在为用户生成的推荐列表中，若目标物品出现在推荐列表中的前 N 个推荐结果内，则视为一次命中[6]。点击命中率计算方法如下：

$$Hit @ N = \frac{Hits}{Sessions} \quad (1)$$

其中， $Hits$ 表示目标物品出现在推荐列表前 N 个推荐结果中被用户交互的会话数量； $Sessions$ 表示测试集中的总会话数。 $Hit @ N$ 指标范围为 0 到 1，指标结果越高说明准确率越高， $Hit @ N$ 可以直观反映会话推荐系统预测的准确率。

2.3. 新颖度

新颖度可以分为项目新颖度和会话推荐系统新颖度。项目新颖度用来衡量在用户与系统交互的过程中，某项目是否频繁出现。若该项目新颖度较高，说明其出现频率较低。会话推荐新颖度则是推荐列表所有物品的平均新颖度，可用于衡量会话推荐系统新颖度，推荐列表中物品的新颖度越高，该系统新颖度越高。

本文物品新颖度基于物品流行度[7]进行计算。物品 i 新颖度 $Novelty_i$ 计算方式如下：

$$Novelty_i = 1 - \frac{click_i - \min(clicks)}{\max(clicks) - \min(clicks)} \quad (2)$$

其中 $click_i$ 代表物品 i 的被点击次数， $\min(clicks)$ 代表数据集中最少被点击的物品的点击次数， $\max(clicks)$ 代表数据集中最多被点击的物品的点击次数。通过上述基于流行度的新颖度计算方式，被点击次数越少(即越小众)的物品具有更高的新颖度。

将所有推荐列表中所有物品平均新颖度值作为会话推荐系统新颖度。若该值较高，则说明系统在保留准确推荐的同时，能够更多地将非热门的物品呈现给用户。会话推荐系统新颖度计算公式[7]如下：

$$Novelty = \frac{1}{U \cdot L} \sum_{u=1}^U \sum_{i \in K_u} Novelty_i \quad (3)$$

其中 U 表示用户数量； L 表示推荐列表长度； K_u 表示用户 u 的推荐列表。

3. 基于图神经网络的新颖化会话推荐系统

3.1. 构建有向会话图

在传统会话推荐算法中，为了反映用户的会话行为先后关系，使算法能够更好捕捉用户行为模式和偏好变化，通常使用线性有序列表作为会话数据的格式。但是线性有序列表只能捕捉相邻两项目的转移关系，而非直接相邻的项目之间无法建立有效的数据关系[8]，线性序列表结构如图 1 所示。在图数据结构中，任意两个节点可以通过边连接，因此能够显式表示所有项目之间的潜在关联[9]。图结构数据如图 2 所示。而图神经网络适用于处理图数据结构，核心是通过邻域聚合生成节点嵌入向量，可以做到同时捕捉局部结构和全局拓扑特征[10]。

基于用户行为序列的会话图构建可定义为：根据给定用户交互行为序列 $s_i = \{v_1, v_2, \dots, v_n\}$ (v_n 表示用

户与系统的交互项), 将每一个会话数据都构建为一个有向图结构 $G = \{N, E\}$, 其中 N 代表有向图的节点, E 代表有向图的边。有向图中的每一个节点则代表一个唯一项目。有向图边的顺序代表用户点击顺序, 边的信息反映了用户与系统交互过程中, 项目跳转的概率。



Figure 1. Data of linear order list

图 1. 线性序列表数据

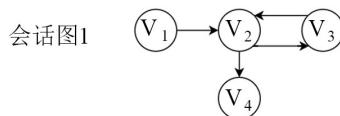


Figure 2. Graph structure data

图 2. 图结构数据

3.2. 项目特征向量计算

为了充分利用 3.1 中提出的图数据信息, 本文引入邻居信息聚合和门控更新机制两种机制。邻居信息聚合指的是通过融合每个节点的邻居信息, 用于捕捉项目之间的转移模式; 门控更新动态能够选择保留原有信息和引入新信息之间的比例, 提高节点状态更新的灵活性。

1) 邻居信息聚合

节点 $v_{s,i}$ 代表会话 s 中的第 i 个项目, 节点 $v_{s,i}$ 在目前时间步 t 中从其邻居节点处整合得到特征向量 $a_{s,i}^t$ 的方法如下:

$$a_{s,i}^t = A_{s,i} \cdot \begin{bmatrix} v_1^{(t-1)} \\ v_2^{(t-1)} \\ \vdots \\ v_n^{(t-1)} \end{bmatrix} + b \quad (4)$$

其中, $v_n^{(t-1)}$ 代表在上一个时间步中, 第 n 个节点的状态向量。 $A_{s,i}$ 是一个维度为 $2n * n$ 的矩阵, 代表节点 $v_{s,i}$ 对应的连接信息向量矩阵, 用于描述该节点与其他节点之间的关系。矩阵 $A_{s,i}$ 由两个部分拼接而成: 第一个矩阵表示节点 $v_{s,i}$ 的出边(从当前节点转移到其他节点的关系), 维度为 $n * n$; 第二个矩阵表示节点 $v_{s,i}$ 的入边(从其他某节点转移到当前节点的关系), 维度为 $n * n$ 。因此, 矩阵 $A_{s,i}$ 中包含节点 $v_{s,i}$ 在图中所有邻接节点的加权信息。 b 代表偏置项, 模型在加权求和后引入偏置项, 补偿由于仅依赖邻居向量加权造成的信息不足, 增强模型的表达能力。

2) 门控更新机制

门控更新机制源于门控循环单元的设计思想, 目标是在信息传播过程中动态确定保留旧信息和引入

新信息之间的比例,从而提高节点状态更新的灵活性和表达能力[11]。在本文中,门控更新机制通过计算更新门、重置门、候选状态和最终状态来实现。

更新门用于控制当前节点状态更新中保留旧状态与采用新候选状态的比例,其计算方式如下:

$$z_{s,i}^t = \sigma(W_z a_{s,i}^t + U_z v_i^{(t-1)}) \quad (5)$$

$a_{s,i}^t$ 表示节点 $v_{s,i}$ 经过邻居信息聚合后,在时间步 t 中所获得的信息,代表当前节点和他的邻域的局部特征相融合的结果。 W_z 和 U_z 代表权重矩阵,分别用于将聚合信息 $a_{s,i}^t$ 和前一个状态 映射到同一特征空间,它们决定了不同信息在计算更新门时的贡献程度。激活函数则将线性组合的结果映射到 $[0, 1]$ 区间,使得每个元素都表示更新比例的权重。

重置门决定了在生成候选状态时,有多少比例的前置状态的成分被重置,从而使得候选状态能更充分地利用当前邻域的信息,重置门计算方式如下:

$$r_{s,i}^t = \sigma(W_r a_{s,i}^t + U_r v_i^{(t-1)}) \quad (6)$$

其中 W_r 和 U_r 是待学习的权重矩阵,分别作用于聚合信息与前一时刻的节点状态,维度匹配保证后续运算正常。

候选状态 \tilde{v}_i^t 表示在当前迭代中根据新信息计算得到的中间状态,反映了当前输入信息对节点状态的一个更新提议。

$$\tilde{v}_i^t = \tanh(W_0 a_{s,i}^t + U_0 (r_{s,i}^t \odot v_i^{(t-1)})) \quad (7)$$

矩阵 U_0 对经过重置门调制的历史状态进行线性变换,与当前邻居信息进行融合。 \tanh 函数将最终线性组合后的结果映射到 $[-1, 1]$ 的区间之内,产生候选状态。候选状态包括融合后新产生的信息和部分历史信息。

最终状态代表节点 v_i^t 在当前时间步 t 中迭代更新后的最终状态,其计算方式如下:

$$v_i^t = (1 - z_{s,i}^t) \odot v_i^{(t-1)} + z_{s,i}^t \odot \tilde{v}_i^t \quad (8)$$

最终状态包含了历史信息与邻居信息,模型不仅能够在信息传播过程中避免过度遗忘历史信息,还可以引入新的上下文信息,从而捕捉会话图中复杂的项目转移模式。

3.3. 生成会话向量

在生成会话向量过程中,本文将局部兴趣和全局偏好信息融合,组成会话的表示向量。局部兴趣用于反映用户实时行为,使用会话中最后一次点击的项目向量来表示;全局偏好用于反应会话中整体的点击模式和隐含偏好,使用聚合所有点击项目的嵌入信息获得[11]。

对于每一个会话,本文使用会话中最后一个项目的嵌入向量局部向量 s_l ,即

$$s_l = v_{s,n} \quad (9)$$

本文认为最后点击的项目恰好可以反映用户当前短期兴趣,因此用来作为短期偏好代表。

使用会话中所有项目向量加权聚合后的结果构造全局会话向量 s_g 。对于会话中每个项目 $v_{s,i}$,注意力权重 a_i 的计算方法如下:

$$a_i = q^T \sigma(W_1 v_{s,n} + W_2 v_{s,i} + c) \quad (10)$$

其中, $v_{s,n}$ 是最后点击项目的嵌入向量,代表局部兴趣; $v_{s,i}$ 是会话 s 中第 i 个项目的嵌入向量; W_1 与 W_2 是模型自动学习的权重矩阵,用于对最后点击和各个项目向量进行线性变换。 c 是模型自动学习的偏置项。

q 是参数向量,用于调整注意力分布,使最终的权重能突出关键项目。在初始化阶段,权重矩阵、偏置项

和参数向量会通过随机赋值以进行初始化，而在模型的每次迭代过程中，损失函数计算出预测结果与真实标签间的损失，再通过反向传播方法计算损失梯度，根据梯度信息更新权重矩阵、偏置项和参数向量的值。

通过比较每个项目与最后点击项目之间的关系，确定各项目对会话全局偏好的贡献度。利用计算得到的注意力权重，对会话中所有项目的嵌入向量进行加权求和，得到全局会话向量 s_g ：

$$s_g = \sum_{i=1}^n a_i v_{s,i} \quad (11)$$

为了充分利用局部与全局的会话信息，本文将局部信息向量与全局会话信息向量拼接，线性变换压缩到同一向量空间，得到最终的会话表示 s_h ：

$$s_h = W_3 [s_l; s_g] \quad (12)$$

其中， W_3 是一个可学习的权重矩阵，使会话表示中包含用户的即时兴趣和整体行为信息。

3.4. 模型训练

为了预测用户与会话的下一次交互，模型为所有项目生成准确性概率，并选取分数排名较高的项目作为推荐预选列表。预选列表的准确性概率使用生成预测概率进行计算，并采用交叉熵损失函数提升预测性能。

1) 计算项目的准确性概率

在获得每个会话的最终表示 s_h 后，模型需要为候选项目生成准确性推荐得分。对于每一个候选项目 v_i ，本文使用项目与会话的相似度进行项目的准确性评估。准确性评分的计算方式为：

$$\hat{z}_i = s_h^T v_i \quad (13)$$

利用会话向量 s_h 与候选项目向量 v_i 之间的相似性来度量项目对当前会话的相关性。内积运算能够捕捉两者在同一嵌入空间内的相似度，得分越高表示该项目越可能是用户下一次点击的目标。最后，为了方便后面推荐得分的计算，将得分映射为正确性推荐概率：

$$\hat{y}_i = \text{softmax}(\hat{z}_i) \quad (14)$$

每个项目预测概率可以被理解为模型认为该项目与用户交互的可能性。

2) 使用损失函数优化推荐准确性

为了优化模型，使预测结果更接近真实点击，本文采用交叉熵损失函数。对每个会话图而言，损失函数定义如下：

$$L(\hat{y}) = -\sum_{i=1}^m [y_i \log \hat{y}_i + (1-y_i) \log (1-\hat{y}_i)] \quad (15)$$

其中 y_i 为真实标签的编码向量，表示实际点击的项目； m 表示候选项目的总数； \hat{y}_i 代表模型预测出的项目 i 被用户交互的概率。

3.5. 新颖性推荐机制

为了在保证推荐准确性的前提下提升推荐结果的新颖性，本节引入一种基于双阶段排序的新颖性推荐机制。该机制的核心思想是在推荐流程中先基于准确性评价筛选出一批高质量候选项，再在候选范围内引入新颖性指标进行加权排序，使最终推荐既具有较高关联度，又能够呈现具有探索价值的内容。与直接对全部候选项进行融合排序相比，这种方式可避免新颖但无关或质量较低的项目进入最终推荐列表，从而有效提升结果质量与用户体验。

为了保证会话列表的准确率，首先选取部分准确性较高的项目生成一个预选集。预选集的生成基于

3.4 小节中提到的会话向量与项目向量的相似性计算。根据每个项目的值，模型按从高到低的顺序对项目进行排序，并选取前 30 个项目构成预选集。预选集的生成确保了推荐结果的准确性和相关性。

对于预选集中的所有项目，按照 2.3 中提到的项目新颖度计算公式计算新颖度，并转换为百分制的新颖度得分 $SNOV_i$ 。将每个项目的准确率预测值转换为百分制的准确率得分 $SACC_i$ 。最终，将新颖度得分 $SNOV_i$ 和准确率得分 $SACC_i$ 结合幂运算后相乘，得到每个项目的最终得分。计算方式下：

$$S_i = (SACC_i)^{(1+x)} * (SNOV_i)^{(1-x)} \quad (16)$$

其中，参数 x 是一个可由用户调节的参数，其取值范围为 $x \in [-1, 1]$ 。综合得分的计算可让系统管理者或用户通过修改参数，对准确度和新颖度进行调制，最终平衡推荐准确性与新颖性。将预选集内所有项目根据最终得分 S_i 从高到低排序，并选取得分排名前 20 的项目作为最终推荐结果。新颖性机制实现的伪代码如下所示：

算法 1：新颖性机制

输入：会话向量 s_h ；候选项目集 I ；参数 x ；

输出：推荐列表 R ；

1. 将 I 中的项目按照 \hat{z}_i 排序，选取前 30 个项目构成预选集 P ；

2. for (i in P) → //循环遍历所有项目

{

$SACC_i = \hat{y}_i * 100$ ； → //将 \hat{y}_i 转换为百分制的准确性得分

$SNOV_i = Novelty_i * 100$ ； → //将 $Novelty_i$ 转换为百分制的新颖性得分

$S_i = (SACC_i)^{(1+x)} * (SNOV_i)^{(1-x)}$ ； → //根据参数 x 计算综合推荐得分

}

3. 按照综合得分 S_i ，选取 P 中项目分数最高的 20 项作为推荐列表 R ；

4. return R ；

4. 实验

4.1. 实验数据集与预处理

本文使用 Yoochoose 和 Diginetica 数据集作为会话推荐模型学习的数据，验证所提出的模型在真实场景下的准确性与新颖性指标。

在数据集中，每一条数据代表了一次真实的用户与系统交互行为，因此会含有无效会话或异常交互。为了让模型能够更好地学习数据集，本文预处理中先对数据进行清洗，再建立会话序列。

1) 数据集分割：为适配实验环境计算规模，本研究将 Yoochoose 数据集缩减至原始规模的 1/128，将 Diginetica 数据集缩减至原始规模的 1/2，后续实验将验证该规模下的模型有效性。在下文中，称缩减后的 Yoochoose 数据集为 Yoochoose 1/128 数据，称缩减后的 Diginetica 数据集为 Diginetica 1/2 数据。

2) 数据读取与会话序列建立：从原始数据集中读取交互记录并创建会话数据。由于原始记录没有严格按照时间顺序存储，为避免生成的会话列表发生顺序上的错误进而误导模型学习，本文将每个会话内的点击记录按照时间戳排序，确保物品序列按照实际发生顺序排列。

3) 会话数据过滤：在数据清洗过程中，为了避免单次点击无法提供有效序列信息，本文将仅包含一个项目的会话删除掉。为了降低数据的噪声，本文删除数据集的所有会话中出现次数少于 5 次的物品。

若某个会话在低频物品过滤后项目数量少于 2，则该会话也被舍弃。

4) 测试集和训练集划分：为了模拟真实的会话推荐系统环境，在评估两个数据集时，选取最近一天的会话数据作为测试集测试会话推荐系统的性能，将其他较旧的会话作为训练集来训练会话推荐模型。

5) 物品重编号：为了方便后续的物品向量生成并降低物品标识的稀疏性，本文将训练集中的物品进行重编号处理，将所有物品标识转换为从 1 开始的连续整数。在测试集中，删除训练集中没有出现过的物品。

4.2. 实验结果

为测试加入新颖性机制的会话推荐系统在准确率与新颖性方面表现，本研究在模型测试阶段模拟了会话推荐系统与用户交互时场景。首先由会话推荐系统学习 Yoochoose 和 Diginetica 数据集训练模型，然后让模型根据综合评分给出 20 个物品作为推荐列表。在评价指标上关注推荐列表的点击命中率和新颖度两方面[12]。

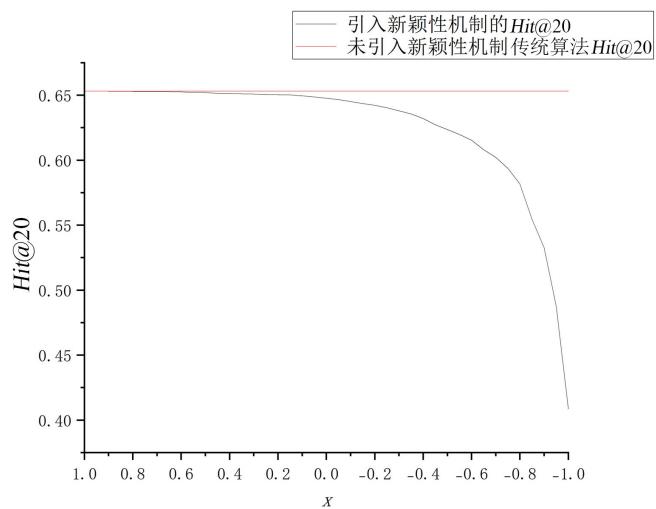


Figure 3. Hit@20 values for different values of x when training with the Yoochoose dataset
图 3. 使用 Yoochoose 数据集训练时取不同 x 值时 Hit@20 值

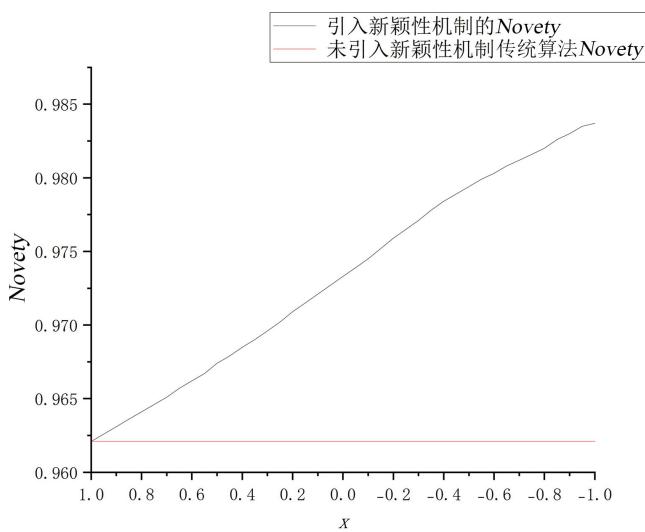


Figure 4. Novelty values for different values of x when training with the Yoochoose dataset
图 4. 使用 Yoochoose 数据集训练时取不同 x 值时 Novelty 值

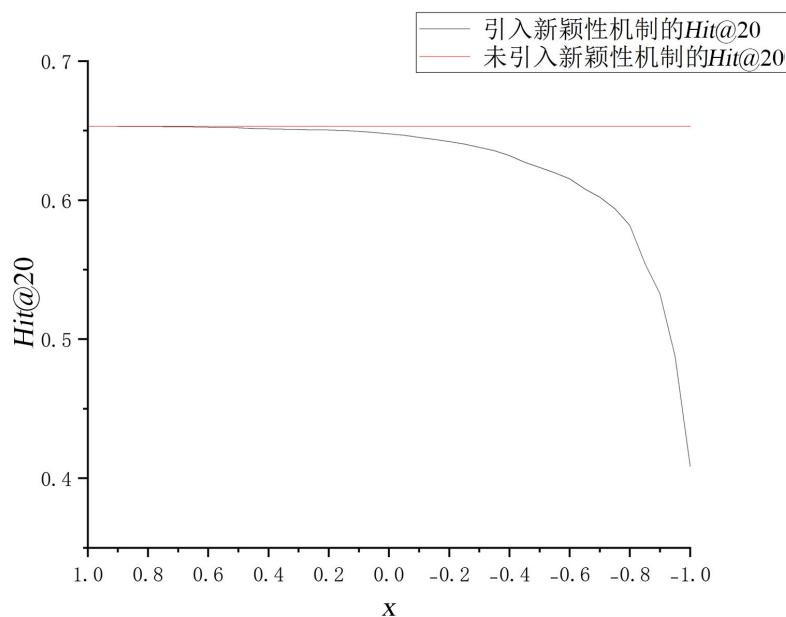


Figure 5. *Hit@20* values for different values of x when training with the Diginetica dataset
图 5. 使用 Diginetica 数据集训练时取不同 x 值时 $Hit@20$ 值

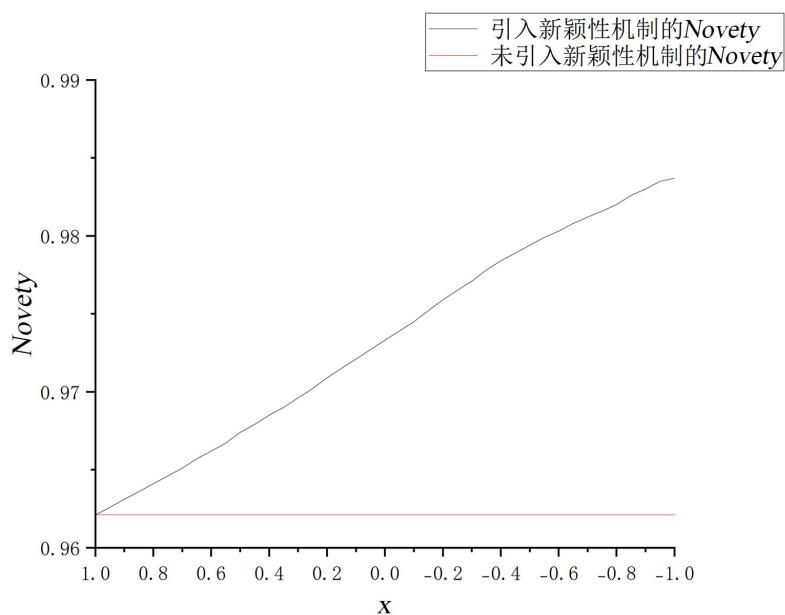


Figure 6. *Novelty* values for different values of x when training with the Diginetica dataset
图 6. 使用 Diginetica 数据集训练时取不同 x 值时 $Novelty$ 值

在使用新颖性算法中参数 x 从-1 到 1 的变化，可以分别得到使用 Yoochoose 和 Diginetica 数据集进行训练的模型的 $Hit@20$ 和 $Novelty$ 两个指标及其变化趋势分别如图 3~6 所示。

分析上面两个数据集训练的推荐列表指标数据，在参数 x 从 1 向-1 降低的过程中 $Hit@20$ 和 $Novelty$ 的变化，将整个新颖性机制推荐的过程分为三个阶段：第一个阶段中相较传统模型的 $Hit@20$ 没有损失，但是新颖性提升较小；第二个阶段中相较传统模型的 $Hit@20$ 损失较小，并且 $Novelty$ 有所提升；第三个阶段中相较传统模型的 $Hit@20$ 损失较大，新颖性有所提升。下面本文针对每个阶段进行分析。

当 x 值为 1 时, 此时模型仅使用准确性得分进行推荐列表筛选, 此时相当于未加入新颖化推荐的传统会话推荐算法。使用 Yoochoose 数据集进行训练的模型 $Hit@20$ 值为 0.6530, $Novelty$ 值为 0.9621; 使用 Diginetica 数据集进行训练的模型 $Hit@20$ 值为 0.4597, $Novelty$ 值为 0.9536。

使用 Yoochoose 数据集进行训练的模型中当 x 从 1 下降到 0.90 的过程中, 推荐系统 $Hit@20$ 仍为 0.6530, $Novelty$ 由 0.9621 提升至 0.9631; 使用 Diginetica 数据集进行训练的模型, 当 x 从 1 下降到 0.9 的过程中, 推荐系统 $Hit@20$ 仍为 0.4597, $Novelty$ 由 0.9536 提升至 0.9547。此时在没有牺牲准确性的情况下, 提升了推荐新颖度。

当 $Hits@20$ 降低 1% 以内, 本文认为相较传统模型 $Hits@20$ 损失较小。在使用 Yoochoose 数据集进行训练的模型中, 当 x 从 0.9 下降到 -0.15 的过程中, 在 $x = -0.15$ 时 $Hits@20$ 数值为 0.6436, 比 $x = 1$ 时的传统算法($Hits@20 = 0.6530$)降低了 0.94%。而 $Novelty$ 却从在 $x = 0.9$ 时的 0.9621 提升到了 $x = -0.15$ 时的 0.9752, 提升了 1.31%。使用 Diginetica 数据集进行训练的模型, 在 x 从 0.9 下降到 -0.05 的过程中, 在 $x = -0.05$ 时的 $Hits@20$ 数值为 0.4499, 比未引入新颖性机制传统算法($Hit@20 = 0.4597$)降低了 0.98%。而 $Novelty$ 从 0.9536 提升到了 $x = -0.05$ 时的 0.9661, 提升了 1.25%。

当准确率降低超过 1% 之后, 本文认为新颖性机制对推荐系统已造成了较大的影响。在使用 Yoochoose 数据集进行训练的模型中, 当 x 从 -0.15 下降到 -1 过程中, x 取 -1 时的 $Hit@20$ 为 0.4086, 比 $x = 0$ 时的值 0.6477 降低了 23.91%; $Novelty$ 从在 $x = 0$ 时的 0.9733 提升到了 $x = -1$ 时的 0.9837。在使用 Diginetica 数据集进行训练的模型中, 当 x 从 -0.05 下降到 -1 过程中, x 值取 -1 时的 $Hit@20$ 为 0.2659, 比 x 取 -0.05 时降低了 18.4%; $Novelty$ 从在 $x = -0.05$ 时的 0.9661 提升到了 x 取 -1 时的 0.9825, 本阶段中虽然新颖度提升了 1.64%, 准确率指标严重下滑, 本文认为这样的准确性指标下滑是无法接受的。因此, 可以认为本阶段的新颖性提升消耗了较大的准确性。

针对上述两个数据集训练的实验结果, 本文根据参数 x 取不同值所对应的 $Hit@20$ 和 $Novelty$, 将推荐系统按照准确率和新颖性结果划分为三个模式: 精准推荐模式、新颖性推荐模式、新颖性优先模式。

精准推荐模式(不牺牲准确率): 精准推荐模式中, 要求相较传统模型的 $Hit@20$ 指标没有损失。例如, 使用 Yoochoose 数据集训练 $x \in [1, 0.9]$ 时和使用 Diginetica 数据集训练 $x \in [1, 0.9]$ 时。在本模式中, 新颖度提升有限, 但是可以保持较高准确性, 既保证了推荐准确性, 又拓展了用户视野。建议在精准投放、时效敏感的业务场景使用。

新颖性推荐模式(推荐准确率降低较少): 新颖性推荐模式中, 要求准确率相较传统模型损失不得超过 1%。例如, 使用 Yoochoose 数据集训练时 $x \in (0.9, -0.15]$ 和使用 Diginetica 数据集训练时 $x \in (0.9, -0.05]$ 。本模式中通过小幅度牺牲准确率, 较大提升了会话推荐的新颖度, 可用于常规推荐场景或稳定交互阶段。

新颖性优先模式(牺牲准确率较高, 新颖性较高): 如果准确率相较传统模型准确率损失超过 1%, 则认为准确性推荐能力损失较大。例如, 使用 Yoochoose 数据集训练时 $x \in (-0.15, -1]$ 和使用 Diginetica 数据集训练时 $x \in (-0.05, -1]$ 。本模式中的新颖性较高, 但是较大地降低了推荐准确性。综上所述, 建议本模式仅在对于新颖性有较高要求的场景下使用, 例如拓展用户视野的冷启动阶段。

系统管理者可根据实际用户使用情况选择上述三种推荐模式, 设置参数以优化会话推荐系统的推荐能力。

参考文献

- [1] 余江龙, 宋腾飞, 汪德超, 等. 个性化推荐系统研究方法综述[J]. 电脑知识与技术, 2024, 20(10): 46-49.
- [2] 齐光鹏. 基于多信息融合分析的客户精准画像与推送算法设计[J]. 现代电子技术, 2025, 48(6): 175-179.
- [3] 翟海民, 魏婷. 人工智能技术在农机产品推荐系统中的应用[J]. 南方农机, 2025, 56(2): 177-179.

-
- [4] 袁凤源, 梅红岩, 温民伟, 等. 基于深度学习的会话推荐方法综述[J]. 辽宁工业大学学报(自然科学版), 2024, 44(1): 6-10+17.
 - [5] 李晶皎, 孙丽梅, 王骄. 提高会话推荐多样性的 SRL 推荐系统模型[J]. 东北大学学报(自然科学版), 2013, 34(5): 650-653+662.
 - [6] 肖楠. 基于会话的推荐算法研究综述[J]. 现代计算机, 2019(36): 33-36.
 - [7] 徐元萍, 陈翔. 推荐系统中的新颖性问题研究[J]. 计算机应用研究, 2020, 37(8): 2310-2314.
 - [8] 张雄涛, 祝娜, 郭玉慧. 基于图神经网络的会话推荐方法综述[J]. 数据分析与知识发现, 2024, 8(2): 1-16.
 - [9] 周芳泉, 成卫青. 基于全局增强图神经网络的序列推荐[J]. 计算机科学, 2022, 49(9): 55-63.
 - [10] 黄震华, 林小龙, 孙圣力, 等. 会话场景下基于特征增强的图神经推荐方法[J]. 计算机学报, 2022, 45(4): 766-780.
 - [11] Wu, Z.H., Pan, S.R., Chen, F.W., et al. (2021) A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems*, **32**, 4-24. <https://doi.org/10.1109/tnnls.2020.2978386>
 - [12] 杨长春, 张毅, 刘昊, 等. 基于全局增强图神经网络的会话推荐方法[J]. 计算机工程与设计, 2024, 45(10): 3089-3095.